

Sequence Listing

## Individual Applicant

-----

Street :  
 City : Athens  
 State : GA  
 Country : USA  
 PostalCode : 30602-7272  
 PhoneNumber : 706-542-0928  
 FaxNumber : \_\_\_\_-\_\_\_\_-\_\_\_\_  
 EmailAddress : wparrott@uga.edu  
 <110> LastName : Parrott  
 <110> FirstName : Wayne  
 <110> MiddleInitial : A  
 <110> Suffix : PhD.

## Individual Applicant

-----

Street :  
 City : Athens  
 State : GA  
 Country : USA  
 PostalCode : 30602-7272  
 PhoneNumber : 706-542-0928  
 FaxNumber : \_\_\_\_-\_\_\_\_-\_\_\_\_  
 EmailAddress :  
 <110> LastName : LaFayette  
 <110> FirstName : Peter  
 <110> MiddleInitial : R  
 <110> Suffix : PhD.

## Individual Applicant

-----

Street : 1907 South Milledge Ave  
 City : Athens  
 State : GA  
 Country : USA  
 PostalCode : 30605-\_\_\_\_  
 PhoneNumber : \_\_\_\_-\_\_\_\_-\_\_\_\_  
 FaxNumber : \_\_\_\_-\_\_\_\_-\_\_\_\_  
 EmailAddress :  
 <110> LastName : Kane  
 <110> FirstName : Patrick  
 <110> MiddleInitial : M  
 <110> Suffix :

## Application Project

-----

<120> Title : Arabitol or Ribitol As Positive Selectable Markers  
 <130> AppFileReference : UGA-855R  
 <140> CurrentAppNumber :  
 <141> CurrentFilingDate : \_\_\_\_-\_\_-\_\_

## Sequence

-----

<213> OrganismName : Escherichia coli  
 <400> PreSequenceString :  
 atgaacgaac aatttacatg gctgcacatc gggttagggt cttttcatcg cgcacatcag 60  
 gcgtgggtatc tacaccgttt gcaggtagtg ggcgataaac gctggagcat tgctgcgggc 120  
 aatattcgta atgatgctga acatgtcgta caggcactca gtgcacagaa aggtcgctat 180  
 gtgctggaaa ccgtcagccc ggaaggggta agcgaatatg aagagatcac ctcaattcag 240

```

aagttgatac cgtggcaggc agatttacaa ccgctgattg ctgaaggggc agatccgaag 300
acaaaagtga ttgctttcac cgtcaccgaa ggcgggtact acctgaatac cagtcacaaa 360
ctggaagtta acaatcctga tttagcggca gatcttaaag ggggatgcaa aacaatttac 420
ggtgttatta cccgtatcct cgaagcgcgt atggcaaata acgccggacc actaaccctg 480
atgaacgaac aatttacatg gctgcacatc ggggtagggt cttttcatcg cgcacatcag 540
gcgtgggtatc tacaccgttt gcaggtgatg ggcgataaac gctggagcat tgctgcgggc 600
aatattcgta atgatgctga acatgtcgtg caggcactca gtgcacagaa aggtcgctat 660
gtgctggaaa ccgtcagccc ggaaggggta agcgaatatg aagagatcac ctcaattcag 720
aagttgatac cgtggcaggc agatttacaa ccgctgattg ctgaaggggc agatccgaag 780
acaaaagtga ttgctttcac cgtcaccgaa ggcgggtact acctgaatac cagtcacaaa 840
ctggaagtta acaatcctga tttagcggca gatcttaaag ggggatgcaa aacaatttac 900
ggtgttatta cccgtatcct cgaagcgcgt atggcaaata acgccggacc actaaccctg 960
ctgaattgctg ataacgtgctg ccataatggt gaacgtttcc atgatggcct ggttgagttt 1020
ctccagctaa ctggcaaaaca ggatgtcatc gactggctga gtacaaatac cacttgcccg 1080
aataccatgg ttgaccgcat tacgcctcgt ccggcagcag aacttcgggc acggatcaag 1140
gctcaaacgg gtattgccga taaagcgcgg gtaatgggctg aaacctttat ccagtgggtc 1200
gtggaagata atttcctgta tgtccgtccg gcactggaga aggtcgggtg cgaactgggtg 1260
gcgtcggtaa tcccctatga agaggcgaag attcgcatc ttaactcttc acacagttgc 1320
atcgccctggg caggtacgtt aatcggtcaa aaatatatcc acgaaagcac aatgaccgat 1380
tttatctatc agattgccga ccgctacgtg acggaagatg tcattccttg cttgggcat 1440
aacggtatcg atttgccaac ctaccgtgat gttgtactca agcgttttac caatccacat 1500
attcaggaca ccaaccaacg cgtgcgtgctg gatggtttct cgaaaattcc ggcatgatt 1560
gccccacac tcgagagatg ctaccagcga ggcgttcgcc cgaatgccac cgccatgtta 1620
cctgcactgt tttacgtatt catggagcag tggcatcacg gcaaaactgcc ctatgaatat 1680
caggatggca tccttgatgc accagctgtc tcgcaatgt tacagtctgc cgatcccgtc 1740
gctgtttatg ccagtataa agcgtgtgtt ggcgatttaa ccgaacgtga agattttgcc 1800
gcgttggtgc gcgaaaaaat cgtgcagctc tacgcgttaa ttaactaa 1848

```

<212> Type : DNA

<211> Length : 1848

SequenceName : SEQ ID NO 1

SequenceDescription : arabitol dehydrogenase

Custom Codon

-----

Sequence Name : SEQ ID NO 1

Sequence

-----

<213> OrganismName : Escherichia coli

<400> PreSequenceString :

```

atcgattgag cagtttgctt cacacggcag ctaaattccc gttcagtgcg tgcaagcaaa 60
cagagactat aaattcgccc tggtaaaagg attatatgat gaatcactct gtgccctcta 120
tgaatactcc ccttaattggc aaagttgcag ctatcactgg cgtgcgtca ggtattggcc 180
tgcaatgtgc aaaaacgctg ctcgatgcag gagcaaaggt agtattgatt gaccgggaag 240
gcgacaaact gcacaagatt gtcgctgagt taggcgaaaa cgcgtacgcy ctgcaactcg 300
atctcttcaa taatcagcaa gtcgataaca tgcggcgga cattatcgaa ctggcggtg 360
ggctggatat ttttcatgcc aatgcaggcg cttatatagg cggcccagtg gctgaagggtg 420
atccagatgt ctgggatcgt gtgttaaatac tgaatataaa tgcggcggtt cgctgtgtcc 480
gtgcagtcct gccgcataatg attgcgcaga ggtcgggcga tataattttt accagttcca 540
tcgcgggcgt cgttccggtt atctgggaac cgatctacac cgcgtccaaa tttgccgttc 600
aggcattcgt acacactacc cgccgccagg tttctcaata tggcgtgcgt gtgggtgcgg 660
tgctgccagg accagtagtc actgcctgc ttgatgactg gccaaaagcc aaaatggaag 720
aagccctggc aaatggtagc ctgatgcaac cgattgaagt ggcggaatca gtattgttta 780
tggtgaccgg ctcgaaaaat gtcaccgtgc gagatttagt gatcctgcct ggcagtgtcg 840
atctgtaagg gcgcaatcat gacaataacc aaaaccgtta ttggtgttga tgtgggatca 900
ggcagtgtcc gcgcgggat ttttgatctc aacggatctc tgctatccca tgccacagaa 960
aaaatcacga ctacgcggcg cagcggaagc cgcgtggaac agtccagcca ggagatctgg 1020
caggcgggtc gttcatgtat tcgtaatgcg ctactctgg cagacgtttg tgcacaaagt 1080
gtggcaggca tcggttttga tgccacctgt tctctgggtg tactggataa aaacggtgat 1140
ccattgcctg tcagcccggg aggagatgca aagcaaaata tcattgtgtg gatggatcac 1200
cgcgccaccg aacaagcaga gcgaatcaat gccactcacc atccggtgct gaactacgtc 1260
ggtggtaaaa tttgcctga aatggaaaca ccgaaaattc tctggctgaa agaaaaatatg 1320
ccagagatct acgaacgtgc cggacaattt ttcgatctgg ccgattttct gacctggcgg 1380
gctaccgggtg atttagcgcg ttcagtatgc actgttacct gtaaatggac gtggctggca 1440

```

```

catgaaaatc gctgggatcc agattatttc cgcaccattg gccttgacaga gttagcggat 1500
gaagatttta ttcgcattgg tcatcatatt gttttccccc gaacaccttg tggaaatggt 1560
ttaacagcac aagccgcggc agagatggga ttactccccg gcacacctgt cgcgttaggg 1620
ttaattgatg ctacgcgtgg tggcatcggt acggttggcg tagaagggtg agcgcgaac 1680
aatctcgctg atgttttcgg cacttcttca tgcaccatgg catctaccac ttctccctcg 1740
tttgtagcgg gtgtctgggg gccgtattac agtgcgatgg ttccagggtc gtggttagtt 1800
gaaggcgggc aaagtgtctg gggagcagct attgaccagc tacttgattt ccatccggct 1860
gttgaagaag ctgcgaaat ggcacaacgt gtgaatcagc ccctccccgt ctggcttgct 1920
gatcgaatcc tcgaaaaaac ggcgcaacca tcagatgctg tcgccctggc gaaagggcta 1980
cacgtggtgc cggaaatttct cggaaatcgc gcccccttcg cagatcctca tgccagagcg 2040
gtaattttgt gcctgggtat ggagcgagat ctggataatt tactcgctt gtatatcgct 2100
ggattatgcg gaattgggta tggctgcgc caaattctcg acgctcaaac agcgcagggg 2160
gtagttagta aaaatatcgt tattagcggc ggtgcccggc agcatccact ggtacgacaa 2220
attctggcag atacctgagg tattccggtc attaccacgc aatgctgcga accggtttta 2280
ttaggctcgg ccattcttgg tgctgtcgcc ggaaatattg caccttctgt tggcgaagcg 2340
atgcaacaat tcacccatgt ggataaatat tattatccgc aagaacgcta tcaatctctt 2400
catcatcgtc gatatgaggc ctataagcag ttgcagcata ctgcaaaatt actcagagac 2460
taattaacca gccgcctgac gctgttttca ggcaatcact aataacgact cactccggta 2520
atatcccgga gtgcattcat ctgcacccta aaaacgaggt ttatatgtcc agaaataata 2580
aacagtgggt gggtttgccg ctgcatctga tatggggata tatcgccatc gcagtattta 2640
tgactggtga tggtttcgaa ctgcgattct tatcgacta tattaaatcg ttaggcttca 2700
cacctgcgga agcctctttt gcctttacgc tctacggcct ggcggtgccc ctttccgctt 2760
gggtttctgg ggtagtagcg gaaatcatca cgccgcaaaa aacctgctg attggttttg 2820
tcctatggtg cgtattccat gttctgtttc tggctcttgg attagacag gcaaactatg 2880
ggttaatcct gctgttttac gggattcgtg gtttagcgta tccgctattt ctctactcat 2940
ttatcgctgt tattattcat aacgtgctga gcgaaaattc cagttctgct ctgggttggt 3000
actggcggtg atattcagta gggatcggtg ttgctggcag ttatattccc agttttacga 3060
taccgattat ggggtgaaatg ggaaccttat ggctggcact ggcgttctgc tttgctggcg 3120
gtgtcattgc catgatctcc ttgcgtcatg ttaaacgcc tggacatatg cataatttaa 3180
ctccccgtga gaagtgttga gaattaagtc gggcagtaac tttactttat accaaccgca 3240
atatttttct ctccagtatt gtgcgcatta taaatacctt atcgttattc ggttttgccg 3300
tcattatgcc aatgatgttt gtggatgaac tgggattcac cacctctgaa tggttgcagg 3360
tttgggcgcc atttttcttt accactatct tctctaatat tttttggggg attgtggcag 3420
aaaaaatggg ctggatgcgt gttattcgct gggttggttg cctcgggatg gcagcatcaa 3480
gttttagcgtt ttactacatg ccgcaatact ttggtcacaa ctactggatg gcaatgattc 3540
cggcgattgc tctgggaact tttgttgcgt catttgtgcc gatggccgct gtcttcccg 3600
cactggaacc aaaacacaaa ggtgctgcaa tctcggttta caacctctct gcgggtatgt 3660
ctaacttctt ggctccggca attgcccgtg tggtattacc gtggttttagc actatcggtg 3720
tggtcattgc ctatacagca ttgtatctat tggcctttgt cctttgcgca ttcattcgcg 3780
ttgagcagcc aggattcagt tctgcgccag tgactgagaa ggcattgaat atctcctgaa 3840
aaacgaaacg catcaggcac tcatcctctc cctcatggga gaggatgatt tcacatcagg 3900
caatagtacg tttgttatcg agataaacgt cctgcacggc gttaatcagt ttcacgccgt 3960
cagccatcga t 3971

```

<212> Type : DNA

<211> Length : 3971

SequenceName : SEQ ID NO 2

SequenceDescription : ribitol operon

#### Feature

-----

Sequence: SEQ ID NO 2:

<221> FeatureKey : misc\_feature

<222> LocationFrom : 96

<222> LocationTo : 848

Other Information : ribitol dehydrogenase coding region

CDSJoin : No

#### Feature

-----

Sequence: SEQ ID NO 2:

<221> FeatureKey : misc\_feature

<222> LocationFrom : 859

<222> LocationTo : 2463

Other Information : ribitol kinase coding region

CDSJoin : No

# Feature

-----

Sequence: SEQ ID NO 2:

<221> FeatureKey : misc\_feature

<222> LocationFrom : 2565

<222> LocationTo : 3839

Other Information : ribitol transporter coding region

CDSJoin : No

# Custom Codon

-----

Sequence Name : ribitol operon

# Sequence

-----

<213> OrganismName : Escherichia coli

<400> PreSequenceString :

|   |     |
|---|-----|
| MMNHSVSPSMN TPLNGKVAAI TGAASGIGLQ CAKTLLDAGA KVVLLIDREGD KLHKIVAELG | 60  |
| ENAYALQLDL FNNQQVDNML ADIIELAGGL DIFHANAGAY IGGPVAEGDP DVWDRVLNLN   | 120 |
| INAAFRVCVRA VLPHMIAQRS GDIIFTSSIA GVVVPIWEPI YTASKFAVQA FVHTTRRQVS  | 180 |
| QYGVVRGAVL PGPVVTALLD DWPKAKMEEA LANGSLMQPI EVAESVLFMV TRSKNVTVRD   | 240 |
| LVILPGSVDL  | 250 |

<212> Type : PRT

<211> Length : 250

SequenceName : SEQ ID NO 3

SequenceDescription : ribitol dehydrogenase

# Sequence

-----

<213> OrganismName : Escherichia coli

<400> PreSequenceString :

|  |     |
|--|-----|
| MTITKTVIGV DVGSGSVRAG IFDLNGSLLS HATEKITTTTR RSGSRVEQSS QEIWQAVCSC | 60  |
| IRNALTADV CAQSVAGIGF DATCSLVVLD KNGDPLPVSP EGDAKQNIIV WMDHRATEQA   | 120 |
| ERINATHHPV LNYVGGKISP EMETPKILWL KENMPEIYER AGQFFDLADF LTWRATGDLA  | 180 |
| RSVCTVTCKW TWLAHENRWD PDYFRTIGLA ELADEDFIRI GHIVSPGTP CGNGLTAQAA   | 240 |
| AEMGLLPGTP VAVGLIDAHA GGIGTVGVEG GALNNLAYVF GTSSCTMAST TSPSFVPGVW  | 300 |
| GPYYSAMVPG LWLVEGGQSA AGAAIDQLLD FHPAVEEARE MAQRVNQPLP VWLADRILEK  | 360 |
| TAQPSDAVAL AKGLHVPEF LGNRAPFADP HARAVICGLG MERDLNLLA LYIAGLCGIG    | 420 |
| YGLRQILDAQ TAQGVVSKNI VISGGAGQHP LVRQILADTC GIPVITTQCC EPVLLGSAIL  | 480 |
| GAVAGNIAPS VGEAMQQFTH VDKYYYPQER YQSLHHRRYE AYKQLQHTAK LLRD        | 534 |

<212> Type : PRT

<211> Length : 534

SequenceName : SEQ ID NO 4

SequenceDescription : ribitol kinase

# Sequence

-----

<213> OrganismName : Escherichia coli

<400> PreSequenceString :

|   |     |
|---|-----|
| MSRNNKQWLG LPLHLIWGYI AIAVFMGTGDG FELAFLSHYI KSLGFTPAAEA SFAFTLYGLA | 60  |
| AALSAWVSGV VAEIITPQKT MLIGFVLWCV FHVLFVLVFL GQANYGLILL FYGIRGLAYP   | 120 |
| LFLYSFIVVI IHNVRSESS SALGWYWAVY SVGIGVAGSY IPSFTIPIMG EMGTLWLALA    | 180 |
| FCFAGGVIAM ISLRHVKTPG HMHNLTPREK FAELSRVAVL LYTNRNIFLS SIVRIINTLS   | 240 |
| LFGFAVIMPM MFVDELGFTT SEWLQVWAAF FTTTIFSNIF WGIVAELKMGW MRVIRWFGCL  | 300 |
| GMAASSLAFY YMPQYFGHNY WMAMIPAIAL GTFVAAFVPM AAVFPALPK HKGAAISVYN    | 360 |
| LSAGMSNFLA PAIAVLLPW FSTIGVVIAY TALYLLAFVL CAFIRVEQPG FSSAPVTEKA    | 420 |
| LNIS  | 424 |

<212> Type : PRT

<211> Length : 424

SequenceName : SEQ ID NO 5

SequenceDescription : ribitol transporter

20250306 11:00:00